A meta-analysis of the effects of instructional interventions on students' mathematics achievement

Jacobse, A.E.; Harskamp, E.G.

A Meta-Analysis of the Effects of Instructional Interventions on Students' Mathematics Achievement

A.E Jacobse
E. Harskamp

Abstract

This study examines the impact of interventions in mathematics education in K-6 classrooms through a systematic review of research literature in the period 2000 - 2010. A meta-analysis of 69 independent effect sizes extracted from 40 primary studies involving a total of 6817 students indicated a statistically significant positive average effect (Cohen's d =.58; SE =. 07) of instructional interventions on mathematics achievement.

Studies that used non-standardized tests as measures of mathematics achievement reported larger effects of instructional interventions than studies that used standardized tests. But, there was no difference in effect found between students of higher or lower mathematics ability, and no difference between direct instruction methods or indirect (guided) instruction methods.

In depth discussion of effectiveness of interventions in mathematical sub-domains is presented. The sub-domains are relevant to the content of international curricula and tests (e.g. Trends in International Mathematics and Science Study) namely: Number Sense, Number Operations, Fractions, Ratio and Percentages, Measurement and Geometry and Word Problem Solving. Studies in each of the sub-domains had a significant average effect size, with studies in Number sense showing a larger average effect and in Measurement and Geometry a smaller effect. The paper finishes with discussion of the restrictions of the meta-analysis and recommendations for future research.

Key words: meta-analysis, instructional interventions, mathematics, achievement

1. Introduction

In the Netherlands, just as in some other Western nations, there is great concern about the level of education in mathematics. International studies show that the mathematics achievement of students in primary and secondary education has declined in the past years. For instance, the international TIMMS (*Trends in International Mathematics and Science Study)* study compared the mathematics and science results of students in grade 4 (10 years of age) and 8 of many nations. The study showed that in some countries mathematics performance increased from 1995 to 2007 However, in a quarter of the countries – including the Netherlands- mathematics scores declined significantly in that period (Mullis, Martin & Foy, 2008). Such findings have led to a renewed interest in effective teaching methods for mathematics and in the conditions in mathematics education that lead to higher levels of learning. Besides being relevant for educational researchers, information on effective mathematics teaching is also of great interest to teacher education. Teachers often do not know which ways of teaching are most effective for success of their students. Students are not always well prepared to take standard mathematics tests. More information is needed on practical ways to improve teaching and to secure effective learning time.

There is a growing convergence among Western nations about mathematics aims and objectives that are marked as most important for the national mathematics curriculum (US Department of Education, 2008). Most nations have ordered the mathematics subject matter in similar content topics or strands and national standards for student performances have been proclaimed and are being implemented. The NCTM (2000) speaks of sub-domains of content that students should master. They span the entire range from Kindergarten through primary education to secondary education. For each grade there are standards for students to reach. The NCTM discerned five main mathematics sub-domains. *Number and Operations* is the sub-domain of understanding numbers, developing meanings of operations and computing fluently. *Algebra* is about concepts and techniques to the representation of quantitative relations and for formalizing patterns, functions, and generalizations. Even young children can use algebraic reasoning as they study numbers and operations and investigate patterns and relations among sets of numbers. *Geometry* is about characteristics of geometric shapes and geometric relationship, as well as the use of visualization and reasoning to solve spatial problems. *Measurement* comprises of understanding attributes, units, systems, and processes of measurement as well as applying the techniques, tools, and formulas to determine measurements. And *Data Analysis and Probability* is essential to being an informed citizen and consumer. Students learn to formulate questions and collect, organize, and display relevant data to answer these questions. Even young children can use methods like graphs or tables to order data and reason about it. The international TIMMS study showed that mathematics teachers and experts greatly agreed that these subject topics which are also taken up in the TIMMS-test cover a great part of the mathematics curriculum in elementary education (Meelissen & Drent, 2008; Mullis, Martin & Foy, 2008).

The five strands discerned by NCTM and TIMMS are also recognized in discussions in the Netherlands on standards for mathematics education and the national mathematics curriculum from primary to secondary education (Expert group mathematics education, 2007). In Dutch standards, Algebra is not a separate strand. The teaching of formula for calculations is included in Operations with numbers, and there is an additional strand for ratio: Fractions, Ratio and Percentages (Experts group Mathematics education, 2008). The Dutch mathematics curriculum follows the curriculum of other Western nations. But the Expert group is very concerned about Dutch students' achievement in Number operations and especially the understanding of fundamental arithmetic concepts and facts and the skills in calculation procedures. Also the attainment in Measurement, the fundamental understanding of measurement concepts and proper use of the measurement system are reason for concern, according to the committee. That is why the Dutch government speeded up the implementation of standards with different levels of mastery for each sub-domain and a basic level for all students. From August 2010 on, all schools for primary education and secondary education are bound to attain the basic level with their students, the national educational inspectorate will monitor schools, projects are started to help schools organize the upgrading of their mathematics education and there is a national support system to help schools and teachers implement standards based mathematics education (Ministry of Education, 2010). The outcome of the results of the Programme for International Student Assessment (PISA) of 15 year old students in 2009 confirmed the government's conclusion that the Dutch students do not longer rank with the top ten of nations and that efforts must be taken to bring up the mathematics level of all students, but especially the more talented students (OECD, 2010, p.19).

When studying mathematics interventions, it is particularly important to assess the effectiveness of instructional approaches in the different sub-domains in mathematics. Different approaches may be effective for different domains and also for different groups of students. The present study examines the impact of instructional interventions on mathematics education through a systematic review of existing literature. This extensive meta-analysis of existing empirical evidence focuses on student mathematics achievement and optimal conditions for mathematics learning in Kindergarten to grade 6 classrooms.

2. Theoretical background

2.1 Mathematics performance in mathematics sub-domains

As discussed above, there is general agreement internationally about the importance of certain mathematical sub-domains. In line with these findings, our study will focus on research literature on instructional approaches in: Number sense, Operations, Fractions, Ratio and Percentages, Measurement and Geometry. These mathematical topics are briefly discussed here to highlight their content and importance for mathematics education.

Number sense. There is general agreement among mathematics educators that children need a well-developed number-sense in order to solve mathematics problems in elementary school. Jordan, Kaplan, Oláh, and Locuniak (2006) illustrated that foundational aspects of number sense include: Counting, number recognition, number knowledge, nonverbal calculation and estimation. Counting is about understanding one to one correspondence of number and object, knowing stable order in counting objects, knowledge of cardinality principal (last counted number indicates total), and knowing the count sequence. Number knowledge is about discriminating quantities and making numerical magnitude comparisons. Estimation of set sizes using numbers as a reference point (e.g. this group of objects is less than 15 but more than 10) becomes part the number sense skills. When a child acquires number sense he or she develops skills in number transformation. Number transformation has to do with transforming sets through addition and subtraction with and without physical objects. Such skills are all at the foundation of being able to perform mathematical tasks. Therefore, most explicit number sense interventions take place in the preschool and the lower grades.

Operations. There are different types of mathematical operations for students to learn during elementary education. Arithmetic addition and subtraction problems to 20 are the basis of operations with whole numbers. In developing competence with these basic facts, children gradually gain efficiency in counting strategies and start to follow more formal procedures for calculation. They discover that the last-number principle reflected in counting is the sum-of principle in addition (e.g. 5+3 = is the number that occurs when counting 3 more from 5). Fluent counting-on knowledge and counting-back knowledge makes addition and subtraction easy to learn. It is the basis for seeing how operations are constructed opens the way for more efficient mathematical behavior. As conceptual knowledge about number grows, children discover that multiplication is a short-hand way for addition and division is an efficient way to repeatedly subtract the same number from a larger number (Fuchs, Fuchs, Compton, Powell, Seethaler & Capizzi, 2006).

Fractions, ratio and percentages. It is well known that teaching and learning fractions, ratio, and percentages in the middle grades is a very complex process. This sub-domain includes all sorts of problems involving ratio. Rational numbers are no longer units that can be added or subtracted but, numbers that express a ratio that can be related to any underlying number of units. That is why a wide variety of physical-empirical situations and representations of mathematical objects are needed to learn ratio problems. Duval (2000) distinguishes physical objects and their ratio (if 4 pupils are given 3 chocolate bars, then 8 students get 6 bars), representation of ratio in schematic representation (e.g. a ratio table) and symbolic representations: 3 : 4 = 6 : 8 (each student gets 3/4 of a bar). Ratio indicates a stable proportion in physical situations. This is the basis for calculation with ratio numbers and understanding how fractions can be added and subtracted by equalizing denominators or multiplication of numerators and denominators. As in all ratio problems, in percentages it is especially important to understand the part-whole schema that underlies problems; problems with unknown whole are difficult to solve (e.g.: Trouser cost 56 euro with discount of 20%; how much was the price before discount?). Learning these concepts is particularly relevant for solving ratio problems.

Measurement and geometry. Measurement of length, area and volume are very important in the elementary school curriculum. In the lower grade some basic concepts behind measurement are introduced and in the upper elementary grades, students move on to more complex geometrical concepts like calculating area and volume. Finding the areas of polygonal figures, needs to be taught through the processes of decomposition and re-composition as well as through geometric calculation (Bonotto, 2003). For example, for developing the formulas to measure the area of a polygon with sides of which the lengths cannot be directly determined, students need to: (a) examine how the polygon can be cut up and rearranged into a configuration to which previously learned formulas can be applied; (b) determine the various shapes embedded in a polygon or the area of a polygon circumscribed within a figure and subtracting that value from the total area of the figure; and (c) applying the necessary formulas. The better conceptualization of geometric motions (e.g., recognizing congruent shapes in different positions) a child develops, the better the understanding of 2-D geometry patterns the child can obtain. Geometric concepts in turn facilitate children's connections between the geometry and numerical domains (i.e., counting and later calculating the units covered on a compound figure). In elementary education spatial perspective taking, map reading and the relationship between 2D representations of 3D situations are also part of geometry (Chazan en Lehrer, 1998).

Word problems. This topic is included in the meta-analysis because much of today's teaching and testing in mathematics makes use of more or less realistic contexts in which mathematics problems are placed. Sometimes the problem situation is straight forward (e.g., Peter has 24 car toys and gives 2/3 of the toys to his younger nephew. How many toys does he give away?). But, other problem situations are more complex and students often have difficulty understanding the problem situation. Students need to be taught how to approach word problems. In mathematics curricula there often is no special attention to problem solving. Teachers often think that word problems are no more complex than the calculation involved (van Garderen, 2004). However, comprehending word problems requires different types of skills. It consists of selecting the right elements in the situation and putting them together in the right relations, and also choosing the right plan to implement and solve the problem (Schoenfeld, 1992). In word problems knowledge form the above mentioned sub-domains of mathematics can be applied. However, students may need instruction about how to go about solving these mathematical problems.

Besides specific interventions in the aforementioned sub-domains, international studies show that 'opportunity to learn mathematics' is an important factor explaining differences in student achievement among nations. If teachers pay more attention in their lessons to certain sub-domains of mathematics, then their students will do relatively better in that area than in other sub-domains. This explains part of the differences in mathematics achievement between nations (Tornroos, 2005). Additionally, as Hiebert & Grouws (2007) state, differences in opportunity to learn are important to mathematics achievement, but the way students are taught is also of importance. Depending on the aim of mathematics education, direct instruction can be more effective than guided instruction. Direct instruction is an instructional approach where a teacher explicitly teaches students learning strategies by modeling and explaining why, when, and how to use them. For instance: when and how to apply carry-over and borrowing procedures in addition and subtraction of large numbers. Guided instruction (or 'indirect instruction') on the other hand provides students with opportunities to find out and discuss mathematics solutions under a teacher's guidance. The teacher's principal role is that of coach who advises and provides help, and interactions between teacher and students is extensive. Hiebert & Grouws expect that direct instruction is more effective for teaching concepts and skills, but if students have to learn how to apply their skills guided instruction may be more effective. This topic is also taken up into the meta-analysis.

2.2 Previous meta-analyses and research questions for this study

Kroesbergen & Van Luit (2003) published a meta-analysis of 58 studies of mathematics interventions for elementary students with special needs. Interventions in three different sub-domains were selected: Number sense (preparatory mathematics), Number Operations (basic skills), and Word problem solving (problem-solving). The authors found studies on interventions of basic math skills to be the most effective. The authors argue that such instructional approaches may be easier to teach to special needs students than other more complex mathematical skills. Furthermore, a few specific characteristics were found to influence the outcomes of the studies. Studies of a shorter duration were found to be more effective than long interventions, possibly due to their more specific focus. In addition to the duration of the intervention, the particular method of intervention proved important: Direct instruction and self-instruction were found to be more effective than guided instruction. In the special needs sample, interventions involving the use of computer-assisted instruction and peer tutoring showed smaller effects than teacher-led interventions.

Slavin & Smith (2008) reviewed research on the achievement outcomes of three types of approaches to improve elementary mathematics: Mathematics curricula, computer-assisted instruction (CAI), and instructional programs. Studies included use a randomized or matched control group design, had a study duration of at least 12 weeks, and achievement measures not inherent to the experimental treatment. Eighty-seven studies met these criteria, of which 36 used random assignment to treatments. There was limited evidence supporting differences in effect of various mathematics textbooks (curricula). Effects of computer assisted instruction were moderate. The strongest positive effects on student learning were found for instructional programs directed at improving teaching in class. E.g. forms of cooperative learning, classroom management improvement, programs for motivation of students, and supplemental tutoring programs. The review concludes that instructional programs designed to change teaching practices appear to have more promise than those that deal primarily with curriculum or technology alone.

Li & Ma (2010) examined the impact of computer technology on mathematics education in K-12 classrooms through a systematic review of literature. They examined 46 primary studies and found statistically significant positive effects of computer technology on mathematics achievement. Thus in general, studies found higher effects when students learned with computers as opposed to in more traditional settings. This was especially true for low achievers in elementary education. Also differences between direct and guided instruction approaches in the use of computer technology were found, using a more constructivist approach was reported to be beneficial. Regarding the design, studies using non-standardized tests as measures of mathematics achievement reported larger effects than studies using standardized tests. Furthermore, relatively short interventions were found to be most effective as reported in the study of Kroesbergen & Luit.

From these meta-analyses we may conclude that little effect may be expected from differences between curricula. It is in the use of instructional approaches that most differences in learning gain appear. Especially approach that make teaching more interactive and enhance learning time. The meta-analyses mentioned above give some indications about which types of teaching approaches work well in elementary mathematics education. For instance, for special needs students direct instruction is probably most effective when teaching basic math skills (Kroesbergen & van Luit, 2003). Also, tutoring programs are effective for those who lag behind and well structured cooperative learning will help to increase students' active learning time and their performance (Slavin & Smith, 2008).

But, these studies focus on different target groups and different aspects of mathematics education, such as special needs education (Kroesbergen & van Luit), the use of computer assisted instruction in the classroom (Li & Ma) or programs that help to improve teachers' practices (Slavin & Smith) and do not discuss the effectiveness of specific instructional approaches for sub-domains of mathematics. The objective of this meta-analysis is to assess the impact of mathematics instructional approaches on mathematics learning for students in grades K-6 across and within the sub-domains outlined above. The meta-analysis will provide a statistical approach to synthesize empirical evidence retrieved from a sample of studies published in peer reviewed scientific journals from 2000 to 2010. By examining this sample, one takes into account design characteristics as a way to control for the quality of each study. Our present meta-analysis seeks answer the following research questions:

1. Does mathematics learning with explicit instructional interventions impact mathematics achievement of K-6 students, compared to mathematics learning in traditional settings?

2. What study features moderate the effects of instructional interventions on K-6 students' mathematics achievement?

3. What are the optimal conditions for effective mathematics teaching with the interventions in terms of K-6 students' mathematics achievement in key mathematical sub-domains?

3. Method

3.1 Search procedure

A search procedure was executed to find empirical studies of the effectiveness of mathematics interventions for students in elementary education. An intervention is defined as a specific type of instruction executed for a certain period of time aimed at enhancing mathematics performance within one or more sub-domains. In order to find interventions in these sub-domains, a literature search was performed in the following electronic databases: *ERIC*, *PsycInfo* and *PsycArticles*. The search was performed with the keyword *math** added with the specific keywords referring to the sub-domains in our study (Number Sense, Operations, Fractions/ ratio/ percentages, Measurement and geometry and Word Problem Solving). In addition to math* we used the following keywords in the abstract to define our search: *number sense, number concept, operations, addition, subtraction, division, multiplication, arithmetic, fractions, percent / percentage(s), ratio, decimal, proportion, measurement, geometry, problem solving, metacognition / metacognitive, ratio, area,* *parameter, metric, volume*. We used the limiters *peer reviewed*, *school age* and *publication date 2000-2010* to limit our search. Additionally, articles were added using the snowball procedure (using references in articles which had already been found).

3.2 Selection criteria

The initial search in the databases produced over 2000 references. After this initial search, studies were judged based on strict selection criteria. Studies were included in the meta-analysis when they met the following criteria:

- The article is published in a scientific, peer-reviewed journal in the period of 2000 up till the end of 2010
- The study is executed in elementary education (regular or special education) with students between about 4 to 12 years of age.
- The study has a between groups experimental design with a control group and/or an alternative experimental group. Alternatively, within group crossover designs are included.
- Students are randomly assigned to conditions (experimental design) or groups of students are randomly assigned to conditions (quasi-experimental design).
- The study explicitly aims at enhancing students' mathematical performance in a specific sub-domain of mathematics (see 3.1);
- The study reports sufficient quantitative findings on a mathematical outcome measure to compute a standardized effect size.
- The number of students in the study is appropriate for the analyses in the study.
- The intervention is clearly described and specified.
- The study reports findings on a performance outcome measure. Performance outcome measures typically score answers to math problems as being correct or incorrect. No performance measures including procedural components such as working method, labeling, or strategies are included in the meta-analysis.

The attitude of students towards mathematics was not included as a measure of student outcome, because in only very few studies data on attitude of motivation of students were available.

After the first selection of 112 studies many studies were ruled out because they did not fit the criteria mentioned above. For instance: they were not about instructional interventions, they did not use randomization, or they were not executed in educational settings or with a view to improve students mathematical performance. After the first selection 71 studies were admitted and after close reading, 40 studies were admitted for the meta-analysis (see section 3.4 for further information).

3.3 Coding procedure

All 40 studies were coded by two research assistants using a coding scheme with the variables presented in table 1. Studies were coded in continuous deliberation between the coders to ensure agreement. Additionally, when the raters did not reach agreement, differences were resolved by discussion in weekly consultation sessions with the coders and the researchers.

The raters coded the information for each study as presented in table 1. Specific to this meta-analysis is that it was specifically reported for which mathematical domain the intervention was developed. This makes it possible to compare effects over as well as within domains, making findings more specific. Furthermore, characteristics of the treatment as well of the sample receiving the treatment are collected in order to be able to compare if these characteristics of studies moderate the effects.

Table 1. Variables from the coding scheme used for the meta-analysis

Category	**Coded information**
General information	Title
	Publication year
	Name of the authors
	Name of the journal
Characteristics of the treatment	Mathematical sub-domain
	Direct versus indirect instruction
	Setup of the intervention (whole class-, small groups- or individual intervention)
	Textual description of the intervention
	Duration of the intervention (less than 20 hours or 20 hours and more)
Characteristics of the sample	Type of education (regular of special education)
	Number of students
	Grade level (low: kindergarten-grade 3, high: grades 4-6)
	Ability of the sample (low, mixed, high math ability)
Design	Research questions/ aims
	Experimental versus quasi-experimental design
	Researcher-developed or standardized test
Outcomes	Performance measure used in the meta-analysis
	Main outcomes
	Means and standard deviations
	Alternative outcomes (i.e. effect size, correlation)

For all studies coded in the meta-analysis, one performance outcome measure is coded so the amount of outcome measures would not cause unequal weightings across studies (Swanson & Carson, 1996). Requirements for choosing an outcome measure were:

- The measure should be representative for the mathematical sub-domain in which the study is categorized. For instance, for a word problem solving intervention, scores on the word problem outcome measure would be coded.
- The performance measure is a near transfer measure collected temporarily close to the intervention, but does not contain identical tasks to intervention.

- If provided and suitable for the sub-domain, a standardized performance test is used. If this is not provided, the outcome measure should be more or less comparable to outcome measures reported in other studies in that sub-domain.
- The measure should be of a reasonable internal reliability

A small amount of studies were coded in two domains. This was only done when the content of the intervention explicitly addressed multiple domains and when the researchers used separate outcome measures matching the two domains. For instance, some studies explicitly addressed problem solving procedures as well as practice with calculations. In such a case the outcomes on the computation measure would be coded in the domain of operations and the outcomes on the problem solving measure would be coded in the domain of problem solving.

3.4 Sample

While coding the 71 studies, some studies were yet excluded from the analyses: Ten studies were deleted because more in depth observation of the articles showed the design to use non-random assignment of groups or participants; Four studies were excluded because the treatment did not meet our definition of an intervention (for instance only the format of test items were varied in the design); Eleven studies were deleted because the performance measure not only measured performance but also some form of procedural behavior; Two studies wee excluded because the intervention did not take place in school and had not intention of improving students math skills; Two studies were deleted because they were executes in secondary education; And two studies were excluded because they did not provide sufficient quantitative date to calculate an effect size.

Eventually, the sample of studies used in the meta-analysis is 40 primary studies which were found to match the selection criteria mentioned in paragraph 3.2. The studies are listed in the References section.

3.4 Calculation of effect sizes

To make comparisons among studies possible, outcomes of each study were transformed to a general effect measure (Lipsey & Wilson, 2000). Of the performance measure coded for each study, means and standard deviations, or, if not available, other outcome variables were extracted from the coding scheme to compute weighted effect sizes based on the standardized mean difference *Cohen's d*. Using the standardized mean difference, studies with unlike outcome measures are transformed into a comparably scaled outcome format (Hedges & Olkin, 1985). To account for differences in sampling error related to the sample size in different studies, the mean effect size d was weighted by the variance of the sample. Furthermore, confidence intervals of 95% were determined to determine the spread of scores around the mean effect size per study (Lipsey & Wilson, 2000).

When studies with independent groups reported there to be no systematical differences between the experimental groups in the pretest, the mean difference d was calculated using the sample means of the posttests of both groups divided by the within-groups pooled standard deviations. In studies which reported statistically significant differences on the pretest despite of the random allocation of students or groups, the standardized mean difference was calculated by dividing the difference scores by the within groups standard deviation (Borenstein, Hedges, Higgins & Rothstein, 2009). For studies with dependent samples not reporting the correlation between the pre- and the post measure, a correlation of 0.5 was used to determine the variance, and the standard deviations of the posttests were used for standardization of the effect size. Not all articles reported means and standard deviations of the sample; for two articles we used the standardized mean difference which was directly used from the article, two studies reported a correlational measure, and for one study we used an *F*-value for differences between changes to calculate *d*.

All calculations of the comparable effect sizes have been performed in the Comprehensive Meta Analysis computer program version 2.0 (Borenstein, Hedges, Higgins & Rothstein, 2005).

3.5 Statistical analyses

Data Analyses

After computing all weighted effect sizes, effect sizes could be averaged over studies or sets of studies in the Comprehensive Meta Analysis program (Borenstein et al., 2005). Moreover, effect sizes of all primary studies and subgroups within studies could be compared using homogeneity analyses to determine if they differed significantly (see below).

Homogeneity Analysis

To compare effect sizes from different (sets of) studies, we tested the homogeneity of the weighted effects using the *Q* statistic. Cochran's *Q* refers to the weighted sum of squared differences between individual study effects and the pooled effect across studies. Thus, a significant p value of the *Q* statistic would show there to be large variation between the studies because the studies show more variation than would be expected by the standard errors (Hedges and Olkin 1985). The homogeneity statistics (*Q*) were used for (groups of) weighted effect sizes to determine whether the effect sizes varied statistically significantly, that is, whether the findings shared a common effect size in the population. If *Q* was not statistically significant, a fixed-effects model would be adopted for data analysis. If *Q* was statistically significant, a random-effects model would be used.

Subgroup (Moderator) Analysis

Subgroup (moderator) analysis was used to examine differences in the average effects among groups and the moderated effects of characteristics of instructional interventions. To control for confounding variables and to avoid the "fishing-trip" type of data analysis, we used ANOVA within the multiple regression framework as our primary statistical tool. Specifically, based on the ANOVA analogues for categorical data, a weighted least-squares (WLS) multiple regression analysis was performed on effect sizes. Sample sizes were used as the weighting variable. The first analysis aimed to identify study features that accounted for significant unique variance in the findings. This was done individually, feature by feature. For example, all gender-related variables formed a block. The block of gender composition (of the studies) was examined individually for absolute gender effects. As a result, we determined absolute effects of study features.

4. Results

In total, 69 independent effect sizes were extracted from 40 primary studies involving a total of 6817 students to examine the effects of instructional interventions on mathematics achievement. Of the 40 primary studies, 6 studies were executed in special education or a combination of special and regular education. All studies used randomization, either using random assignment of students to experimental/ control conditions or using statistical control for quasi-experimental designs. All studies were articles published in scientific journals between the years 2000 and 2010.

4.1 Overall Effects

Of the 40 primary studies measuring effectiveness of mathematical intervention, the average weighted random effect size *d* was .58 (*SE*=.07) with a 95% confidence interval from .45 to .72. The fixed effect size *d* was .59 (*SE*=.02) with a 95% confidence interval from .54 to .64. The confidence intervals being above zero shows the effect of mathematical instruction to be significant across studies ($p<.01$). Rosenthal and Rosnow (1984) classified effect sizes more than 0.50 *SD* as large, between 0.30 and 0.50 *SD* as moderate, and less than 0.30 *SD* as small. The weighted averages therefore show that overall, mathematics interventions were found to have a medium to large positive effect on mathematics achievement after sample size was controlled.

The effect sizes retrieved from empirical research studies had a minimum of -.27 (*SE*=.12) and a maximum of 2.5 (*SE*=.26). The effect sizes of the various primary studies differed significantly in the sample (Q = 272.45, df =39, p<.01). We therefore adopted a random-effects model for the analyses.

4.2 Study Features that Moderate the Effects of instructional interventions on Mathematics Achievement

This section addresses the issue of what study features moderate the effects of instructional interventions on mathematics achievement. We analyzed mean effect sizes according to (a) the experimental design and (b) the type of outcome measure used in the study to identify factors that significantly moderated the effects of instructional interventions.

As presented in table 2, we did not find differences between studies using random allocation of students (23 studies) and quasi-experimental studies (17 studies). We emphasize that effect sizes were weighted by sample sizes through the use of weighted least squares (WLS) regression. However, the outcome measure researchers used to determine the effects of their intervention did significantly affect the average effect size. The average effect size found in the 31 studies using a researcher developed test is .64 (*SE*=.07) while the effect on standardized tests is significantly lower (*d*=.32, *SE*=.13).

Table 2. **Mean effect sizes and confidence intervals grouped according to study features**				
	Mean Effect Size (*SE*)	-95% CI	+95% CI	*Q* statistic (df), p value
Experimental	.57 (.07)	.44	.69	
Quasi-experimental	.58 (.13)	.32	.83	.01(1), p=.95
Researcher developed test	.65 (.07)	.51	.79	
Standardized test	.32 (.13)	.07	.58	4.74(1), p=.03

4.3 Effects of sample characteristics

We paid particular attention to the sample characteristics of the studies, coding (a) the grade students were in, (b) students' ability and (c) special education or regular education. We investigated whether these sample characteristics would explain the variance in effect size measures by examining in a separate manner whether each block of variables was related to the effects of technology on mathematics achievement.

As shown in table 3, we found no significant moderating effects on the average effectiveness of mathematics interventions. On average, interventions in special education were found to have a higher effect than interventions in regular education. However, the differences between effects in regular and special education were not significant.

Table 3. **Mean effect sizes and confidence intervals grouped according to sample characteristics**				
	Mean Effect Size (*SE*)	-95% CI	+95% CI	*Q* statistic (df), p value
Grades 0 – 3	.58 (.10)	.39	.76	
Grades 4-6	.58 (.09)	.40	.77	1.81(2), p=.40
Low ability	.54 (.08)	.38	.69	
Mixed ability	.59 (.10)	.40	.78	.17(1), p=.68
Regular education	.55 (.08)	.40	.71	
Special education (and mixed)	.75 (.17)	.43	1.07	1.13(1), p=.29

4.4 Effects of intervention-related characteristics

To examine effects of instructional methods with different characteristics, we coded several implementation features: (a) method of teaching and (b) set-up of the instruction and (c) duration of the intervention (or treatment).

Similar to the case of sample characteristics, we investigated whether these general characteristics of interventions would explain the variance in effect size measures by examining in a separate manner whether each block of variables was related to the effects of technology on mathematics achievement. Table 4 shows that we found no significant intervention-related characteristics which moderated effects on mathematics achievement over all studies.

In general, we found no indication that indirect instruction (following the constructivist approach of guiding students instead of leading them) resulted in better student achievement than direct instruction. We could not investigate (due to the limited number of studies) whether direct instruction is profitable for low math ability students. But we come back to that point of discussion when presenting the outcomes of studies in the sub-domain of Operations on numbers.

On average, small groups interventions or individual interventions taken together yielded higher effects than whole class interventions. But there was no difference in effect for short-term versus long term interventions. However, note that these results were averaged over interventions with differential focus and content.

Table 4.
Mean effect sizes and confidence intervals grouped according to intervention characteristics

	Mean Effect Size (*SE*)	-95% CI	+95% CI	*Q* statistic (df), p value
Direct instruction	.58 (.10)	.39	.77	
Indirect instruction	.61 (.12)	.37	.86	1.82(2), *p*=.40
Whole class intervention	.51 (.15)	.22	.79	
Small groups intervention	.63 (.13)	.38	.88	
Individual intervention	.60 (.10)	.41	.79	.45(2), *p*=.80
Short intervention	.62 (.08)	.47	.77	
Long intervention	.46 (.14)	.18	.75	.90(1), *p*=.34

4.5 Relative of the interventions on Mathematics Achievement

Although assessing overall effects of mathematics interventions is informative to discern how students' math achievement can be stimulated in general, within mathematics the content between sub-domains differs considerably. To get a clear view about what works for different types of mathematical content, we discuss effectiveness and characteristics of interventions specific to the domains of Number Sense, Operations, Fractions, Ratio and

Percentages; Measurement and Geometry; and Word Problem Solving. Within these domains we also took into account differential effects for subgroups (for instance ability groups within the studies) as well as multiple experimental conditions. In the case individual studies had more than two groups, the different groups are compared pair-wise and reported separately. This provides the opportunity to compare different effects reported in individual studies more in depth than in the overall analyses.

Firstly, the average effect sizes within the different domains are presented in table 5. These show that on average, studies in all different domains report medium to large effects of interventions. The lowest effects were reported in the complex domain of measurement and geometry, while number sense interventions relatively had the largest mean effects. Across all domains, no significant differences in effect sizes were found ($Q_{(4)}$=1.85, p=.76)

Table 5.
Mean effect sizes and confidence intervals for the five mathematics sub-domains

	Mean Effect Size (*SE*)	-95% CI	+95% CI
Number Sense	.67 (.11)	.46	.88
Operations	.59 (.11)	.38	.80
Fractions, Ratio and Percentages	.63 (.19)	.25	1.00
Measurement and Geometry	.48 (.13	.22	.74
Word Problem Solving	.50 (.12)	.28	.73

4.5.1 Number sense

Table 6 shows eight studies included in the meta-analysis for the sub-domain of number sense. The general effect size of the experimental condition in these studies is .67 (SE .11). This is an above average effect size in relation to the overall effect size found in all studies.

The studies in this sub-domain can be divided into two groups. The first group with studies by Arnold, Fisher, Doctoroff and Dobbs (2002), Clements and Sarama (2007 and 2008), Starkey, Klein and Wakeley (2004), and Van Luit en Schopman (2000) deals with extensive treatments of number sense. The interventions took place during 6 weeks (Arnold et al.) up to 26 weeks (Clements & Samara). Students were trained every day for a brief period in different aspects of number sense. The activities entail: Counting objects and counting on or backwards, comparison of quantities (how many more or less than ….), comparison of numbers, change operations (addition and subtraction with quantities) and measurement of length, weight and volume (with natural measures). Children's number sense is built up through these activities. Sometimes geometry activities with shapes, patterns and construction in 3D from 2D drawings are also included. The activities are brought to children through books, songs, computer games, projects, role play or group discussions. The programs use a combination of these methods. They represent the various ways teachers are used to create opportunities for young children to learn and play. To test the effect of the programs many broad spectrum of subtests are used, tapping different aspects of number sense.

Table 6.
Overview of studies and standardized effect sizes in the domain of Number Sense

	Subgroup	Mean Effect Size (*SE*)	-95% CI	+95% CI
Arnold Fisher (2002)		.44 (.19)	.06	.81
Clements Samara (2007)		.64 (.25)	.15	1.13
Clements Samara (2008)	Exp 1: Building Blocks	1.09 (.15)	.79	1.38
	Exp 2: Comparison group	.60 (.18)	.26	.95
Ramani Siegler (2008)		.12 (.19)	-.24	.47
Siegler Ramani (2008)		.99 (.35)	.30	1.68
Starkey Klein (2004)		.84 (.16)	.52	1.16
Van Luit Schopman (2000)		.71 (.18)	.34	1.07
Wilson Dehaene (2009)		.74 (.28)	.18	1.29

Note: In the table, the first two authors are mentioned; all authors are added in the reference list

The most effective program is the Building Blocks intervention which is an extensive program developed for pre-kindergarten to grade 2 funded by the National Science Foundation. Control groups in the number sense intervention studies typically continued the program kindergartens offered and did not receive systematic education in pre-school mathematics. In comparison, the programs offered a systematic approach to learning early mathematic skills.

The second group has the studies of Ramani and Siegler (2008), Siegler and Ramani (2008), and Wilson, Dehaene, Dubois and Fayol (2009). In these studies the intervention is brief and the tests are mainly constrained to number identification and comparison. The studies of Siegler and Ramani use a simple board game in which children have to count, on a number line of squares. The game is organized four times and takes 25 minutes on average. In the study of Wilson and Dehaene students are given a computer game and have to count on or backwards on a number line of squares. They play against the computer and the level of the game (speed and complexity of mathematics tasks) is adjusted to the responses of the child. The game was tested in schools and improved children’s number recognition and numerical comparison skills, just as the game of Siegler and Ramani did. The fact that effects of the studies of Wilson et al. and Siegler and Ramani are quite comparable on specific sub-tests of number sense to effects of the larger interventions, shows that number sense can be trained in various, and also time-efficient manners.

4.5.2 Number operations

Table 7.

Overview of studies and standardized effect sizes in the domain of Number Operations

	Subgroup	Mean Effect Size (*SE*)	-95% CI	+95% CI
Booth Siegler (2008)	Exp 1: Child generate	.20 (.28)	-.34	.74
	Exp 2: Computer generate	.75 (.29)	.19	1.31
	Exp 3: Child + computer generate	.20 (.27)	-.34	.73
Fuchs Fuchs (2002)	Low achievers	.25 (.22)	-.19	.68
	Average achievers	.19 (.16)	-.13	.51
	High achievers	.13 (.21)	-.27	.54
Fuchs Powell (2009)	Exp 1: Word Problem tutoring	.62 (.21)	.20	1.04
	Exp 2: NC tutoring	.55 (.21)	.14	.96
Fuchs Powell (2010)	Exp 1: Strategic counting and practice	.67 (.20)	.28	1.07
	Exp 2: Strategic counting	.43 (20)	.03	.83
Kroesbergen van Luit (2002)	Regular education, Exp 1: Guided	.63 (.33)	-.02	1.29
	Regular education, Exp 2: Structured	-.10 (40)	-.88	.67
	Special education, Exp 1: Guided	3.34 (.76)	1.86	4.82
	Special education, Exp 2: Structured	-.55 (.50)	-1.53	.43
Kroesbergen van Luit (2004)	Exp 1: Constructivist	.38 (.15)	.09	.69
	Exp 2: Explicit	.43 (.15)	.13	.73
Nunes Bryant (2009)	Exp 1: Visual demonstration	.32 (.32)	-.29	.93
	Exp 2: Oral Calculator	.19 (,32)	-.44	.82
Ploger Hecht (2009)		.31 (.14)	.02	.59
Schoppek Tulis (2010)		2,5 (,26)	2.00	3.00
Timmermans Lieshout (2007)	Boys	-.26 (.46)	-1.17	.64
	Girls	.62 (.45)	-.26	1.49
Tournaki (2003)	LD, Exp 1: Strategy instruction	2.33 (.49)	1.37	3.29
	LD, exp 2: Drill and practice	.49 (.38)	-.26	1.24
	Regular education, Exp 1: Strategy instruction	1.53(.43)	.69	2.37
	Regular education, exp 2: Drill and practice	1.35 (.42)	.53	2.17
Wong Evans (2007)		.50 (.26)	-.02	1.02
Woodward (2006)		.50 (.05)	.41	.59

Note: In the table, the first two authors are mentioned; all authors are added in the reference list

The sub-domain of Operations is about addition, subtraction, multiplication and division with whole numbers. Our survey of papers yielded 13 studies and 28 effect sizes when reporting on subgroups and different experimental conditions. Studies in the domain of operations on average show a medium effect of interventions on mathematics achievement of students: *d*=.50 (*SE*=.05). The effect sizes of the individual studies are reported in table 7.

The intervention studies can be subdivided into two groups: individualized practice studies, and studies on group teaching.

The individualized practice studies are: Booth and Siegler (2008) who researched the effect of representing numbers with colored bars on a number line and its effect on addition skill; Fuchs et al. (2009 and 2010) and Wong and Evans (2007) who used strategic counting as a backup to improve students' knowledge of number facts to 20. Nunes et al. (2009) investigated the effect of visual demonstration of the inverse actions of addition and subtraction to 20 with unifix bricks hidden under a cloth; Schoppek and Tulis (2010) showed that a training program with visual supports that can adapt the level of practice to the individual student's needs improves the multiplication skill of students and Woodward (2006) established that deliberate practice with a multiplication grid and feedback on errors, enhanced students multiplication skills considerably.

The studies on group teaching are: Fuchs, Fuchs Yazdian and Powell (2002) who performed a study on peer assisted learning strategies in which practice with number lines, sticks (tens), beans (units) and games were provided and students played the role of teacher and student alternatively to learn the number facts and number to 100; Kroesbergen en van Luit (2002 and 2004) concluded that guided instruction in multiplication teaching with number lines is less effective in special education but more in regular education; Timmermans van Lieshout and Verhoeven (2003) established that guided instruction in addition and subtraction with the number line is profitable for female students with learning disabilities but not for male students. Male students profit more form direct instruction; Ploger & Hecht (2009) investigated the effect of teachers' use of a program to visualize and conceptualize for their students the inverse operations of multiplication and division; Toumaki (2003) made clear that instruction in the use of a counting frame with two rows of 2 x 5 beads is more profitable to students' addition and subtraction skills to 20 than the use of traditional counting materials.

The general theme in the studies discussed above is that conceptual knowledge and skills of students are best supported with the help of visual aids and by training cognitive strategies. This can be achieved by teaching but also with the help of computer programs that are added to teachers' instruction of number operations.

4.5.3 Fractions, ratio and percentages

The sub-domain of fractions, ratio and percentages is about the use of rational numbers. Rational numbers are numbers that express a ratio that can be related to any underlying number of units. For instance 3/4 can represent any number of units (object) en so does 75% or '3 out of 4'. Rational numbers are relative numbers and yet they seem to be operated upon like natural numbers: we can add, subtract, divide or multiply rational numbers. But the rules are different from those for whole numbers. Many students find this hard to understand. In the studies reported here much attention is paid to visual models representing rational numbers in proportion, fractions, percentages and decimal numbers. The assumption is that different visual models for a wide variety of problem situations help to make ratio more tangible and easier to understand. This assumption seems true, on average interventions in the domain of fractions, ratio and percentages have a large effect of d=.63 (SE=.19).

Table 8.
Overview of *s*tudies and standardized effect sizes in the domain of Fractions, Ratio and Percentages

	Subgroup	Mean Effect Size (*SE*)	-95% CI	+95% CI
Cramer Post (2002)		1.00 (.05)	.89	1.10
Huang Liu (2008)		.41 (.15)	.12	71
Tajika Nakatsu (2007)	Exp 1: Self-explanation	1.17 (.30)	.59	1.75
	Exp 2: Self-learning	.14 (.28)	-.41	.68
Van Dijk Van Oers (2003)		.41 (.13)	.25	1.00

Note: In the table, the first two authors are mentioned; all authors are added in the reference list

Cramer, Post & delMas (2002) investigated the development of conceptual knowledge of fractions with the help of circles, chips, bars and number lines as visual models in grade 4 and 5. Fractions were compared, ordered and outcomes were estimated and then calculated. This approach was compared to a standard approach with much stress on declarative and procedural knowledge of fractions. Huang, Liu and Shiu (2008) had a computer program that taught students decimal fractions with the help of cognitive conflict in grade 6. Students were given a multiple choice problem. If they gave an incorrect answer this was confronted with a question to explain their answer. Then instruction was provided and a new problem was offered to make sure the student had understood the instruction. In the instruction several visual models were used (objects, place value schema and fraction circles). Natural fractions were used to explain decimals. Tajika et al. (2007) studied the improvement of skills to solve ratio problems in grade 7. Much use was made of the part-whole scheme. For instance tap A takes 10 minutes to fill a tank and tap B takes 15 minutes. How much time to fill the tank if both taps are open? In a worked example, a schema was used showing that per minute 1/10 of the tank and 1/15 of the tank were filled and the student was asked how much this is together and how long it would take to fill the tank.

Van Dijk & van Oers (2003) gave students a series of lessons in which they guided students through worksheets with problems and group discussion into the use of different visual models for solving problems on percentages. The learning gain as measured by the post test was higher in the experimental group than in the control group in which the teacher directly instructed students the use of visual models for solving problems on percentages.

In the studies visual models play an important role. It turns out that variation in problem situations, visual models that integrate fractions and percentage, decimals or proportion and worked examples are important ingredients for effective teaching of ratio. There is evidence from one study regular students learn to apply visual models in word problems better that through guided instruction than through direct instruction.

4.5.4 Measurement and geometry

Students in elementary school only gradually understand the principles underlying measurement. They first need to develop understanding of measurement in different units and how these units can be ordered according to a (metric) system. For area measurement they need to understand the relationship between the dimensions of the rectangle in covering an area, the measurement outcome and the role of multiplication in calculating area. Much research has been done on the development of children's understanding of area measurement (Bonotto, 2003; Kamii & Kysh, 2006) and volume (Battista, 2003). Researchers agree that students need to link their intuitive idea of measuring by counting units to multiplication with the help of formula. In the past ten years, many observational studies have been done on how to help students (e.g. Iszak, 2005), but few experiments have been undertaken. However, these studies show that teaching students about this sub-domain can affect their performance. The average effect size of the studies in table 9 is of medium level (*d*=.48, *SE*=.13).

Table 9. **Overview of studies and standardized effect sizes in the domain of Measurement and Geometry**

	Subgroup	Mean Effect Size (*SE*)	-95% CI	+95% CI
Hannafin Truxaw (2008)	High spatial	-.05 (.35)	-.74	.64
	Low spatial	.66 (.36)	-.05	1.36
Hung Witz (2011)	Exp 1: Numerical + conceptual instruction	.71 (.27)	.18	1.23
	Exp 2: Numerical instruction	.35 (.26)	-.16	.86
	Exp 3: Conceptual instruction	.43 (.26)	-.08	.94
Olkun Altun (2005)		.71 (.21)	.29	1.12
Souvegnier Kronenberger (2007)	Exp 1: Jigsaw + peer questioning	.20 (.17)	-.14	.53
	Exp 2: Jigsaw	.15 (.17)	-.18	.49
Steen Brooks (2006)		1.80 (.43)	.97	2.64

Note: In the table, the first two authors are mentioned, all authors are added in the reference list

Hannafin, Truxaw, Vermillion, and Liu (2008) investigated the effect of the computer program Sketchpad (with a booklet of activities) versus an online tutorial with written materials on geometric shapes. Sketchpad offered problems for drawing and manipulating of shapes in order to discover their properties (e.g. the properties of a triangle or a rectangle) and the students using the tutorial did activities with the same problems but without software to construct and manipulate shapes. The study is carried out in grade 6 and students in both conditions worked in dyads. The computer program was shown to affect the performance of students with low spatial ability, but not of high spatial students.

Huang & Witz (2011) studied the effect of three conditions of teaching: Numerical calculation, conceptual understanding combined with numerical calculation and conceptual understanding on conceptual understanding and calculation of rectangles and triangles area and perimeter in grade 4. The control group did exercises in multiplication and division with numbers of three digits. In the conceptual condition the row-and-column structure of a rectangular array was related to the properties of basic shapes, concepts of congruence, and geometric motions. In the calculation condition geometric shapes were related to exploring the formulas for area measurement and numerical calculations were included in the problem-solving activities in this condition. In the conceptual with calculation condition both activities were combined. In all conditions students improved their understanding skills in area measurement more than in the control condition. But students in the combined conceptual and calculation condition had the highest gain score.

Olkun, Altun and Smith (2005) used a computer program in grade 4 and 5 with 40 Tangram puzzles that had different geometrical shapes. The shapes had to be used to construct figures form simple to complex. In order to create the figures students had to transform the shapes (rotate, shift, spin etc). The students were asked to compare the area between figures. Students in the control group continued their regular mathematics classes and used no treatment materials. The use of the Tangram puzzles had considerable effect on students' spatial and area measurement skills.

Souvingnier & Kronenberger tried to establish the effect of collaboration in learning about geometric solids (cubes, pyramids, cones and cylinders) in grade 3. Classes were assigned randomly to three conditions: teacher guided practice, co-operation organized in jigsaw and jigsaw with peer questioning. Students did several activities: different perspectives on cuboids, constructing pyramid with cubes, wrapping cones, shadows of cylinders, etc. It turned out that the students in the teacher led condition did not do better in the post tests on geometrical shapes, symmetry and topology than students in the two jigsaw conditions. The researchers saw this as a positive outcome, meaning that co-operative learning can be introduced to vary teacher led lesson in geometry.

Steen & Brooks (2006) investigated the effect of virtual manipulatives on first graders during a geometry unit. The experimental group used tool software to carry out the tasks in a series of worksheets. They had to identify 3D and 2D geometrical shapes, draw plane shapes, state a rule for a pattern of shapes, use symmetry in shapes and show equal parts in shapes. The control group did the same tasks with the worksheets. They could share the use physical 3D and 2D shapes with a group of students and had no computer software. The experimental group outperformed the control group on geometry tests from the mathematics textbook. The researchers assume that the effect of the use of virtual manipulatives is caused by the flexibility of these tools. Recognizing shapes and manipulating them is easier when a 3D figure can be changed into an array of 2D shapes and vice versa, and also when 3D shapes can be taken apart or rotated.

Summarizing, all studies we found in the area of measurement and geometry used tools (squares, grids or adaptable virtual shapes) to bridge the gap between counting units and to measure area or to bridge the gap between concrete manipulation of shapes and knowing the rules to construct different geometrical shapes. This is best done in the primary grades when children are first exposed to measurement. In area measurement they can learn to use tools like the row-column structure in simple rectangles as the basis for calculation of area and in geometry the tools for making patterns and applying rules to build up shapes. It is of importance that students have opportunities to use the tools for themselves.

4.5.5 Word problem Solving

In contemporary mathematics education mathematical tasks are often embedded in a realistic context to make problem solving more authentic and meaningful (Heuvel - Panhuizen, 1996; Kroesbergen & van Luit 2002). However, to solve complex word problems involves more skills than mathematical calculations (Fuchs, Fuchs, Stuebing, Fletcher, Hamlett & Lambert, 2008b; Mayer & Hegarty, 1996; Verschaffel, De Corte, Lasure, Van Vaerenbergh, Bogaerts & Ratinckx, 1999). Students also need skills to analyze the problem, translate information from the problem into a mental model, and use this to find an appropriate solution method (Schoenfeld, 1992). To support students' problem solving processes, many intervention studies have already been undertaken. The experimental studies found in our literature search show that problem solving interventions can have a substantial impact on students' word problem solving performance. The mean effect size d=.50 (SE=.12). The effects of individual studies are presented in table 10.

In general, all studies in this domain try to teach students problem solving steps like the ones Schoenfeld (1992) described. But there are some differences in the focus of the treatments. About half of the studies on problem solving found in this meta-analysis primarily focus on the first phases of the process: analyzing the problem and constructing a mental model or schema of the problem situation. Students are taught how to recognize and analyze different types of problem situations and find a formula to solve the problem. The studies are often restricted to a sub-domain: e.g. recognizing word problems that need an addition or subtraction formula to solve.

In the studies of Fuchs and colleagues (Fuchs et al., 2008; 2010) and Fuchs, Powell, Seethaler and Cirino (2009; 2010) students problem solving process is supported by schema-based instruction. The approach has been proven effective in several studies in elementary education, although some studies also use process measures of the use the schema-based approach as dependent variables. However, in his meta-analysis we only included studies using achievement tests. As shown in table 10, when using achievement tests to measure effect, only medium effect sizes were found on students' word problem solving performance.

Table 10.
Overview of studies and standardized effect sizes in the domain of Word Problem Solving

	Subgroup	Mean Effect Size (*SE*)	-95% CI	+95% CI
Chang Sung (2006)		.78 (.30)	.20	1.36
Chung Tam (2005)	Exp 1: Worked examples	1.40 (.50)	.42	2.38
	Exp 2: Cognitive strategy	1.17 (.48)	.22	2.12
Desoete Roeyers (2003)	Exp 1: Metacognitive training	1.14 (.21)	.73	1.55
	Exp 2: Cognitve training	.44 (.20)	.06	.83
	Exp 3: Motivational training	.16 (.21)	-.26	.57
	Exp 4: Quantative-relational training	.43 (.21)	.03	.84
Fuchs Seethaler (2008)		.35 (.34)	-.32	1.02
Fuchs Zumeta (2010)		-.27 (.12)	-.51	-.03
Fuchs Powell (2009)	Exp 1: Word problem tutoring	.14 (.21)	-.27	.55
	Exp 2: NC tutoring	-.10 (.21)	-.50	.30
Fuchs Powell (2010)	Exp 1: Strategic counting + practice	.36 (.20)	-.03	.75
	Exp 2: Strategic counting	.55 (.20)	.16	.95
Hohn Frey (2002)	Grade 3	.93 (.38)	.18	1.67
	Grade 4	1.06 (.35)	.37	1.75
	Grade 5	.65 (.35)	-.04	1.33
Jacobse Harskamp (2009)		.50 (.30)	-.08	1.08
Teong (2003)		.60 (.32)	-.03	1.24

Note: In the table, the first two authors are mentioned, all authors are added in the reference list

In the other set of studies a more general problem solving approach was offered to students. In the study of DeSoete Roeyers & Clerck (2003) students were first prompted to monitor and predict if they could solve problems and then they tried to solve them. During the solution process, they were supported to execute specific problem solving steps. In the same vein, Hohn and Frey used the ‘SOLVED’ checklist which stands for State the problem, Options to use, Links to the past, Visual aid, Execute your answer, and Do check back. The teachers modeled how to use the checklist and students used this as an aid to solve word problems. Teong (2003) used the so called ‘CRIME’ strategy Careful Reading; Recall Possible Strategies; Implement Possible Strategies; Monitor, and Evaluation to support low ability students’ problem solving in a computer environment. Students’ got a checklist with questions guiding their problem solving behavior such as “Do I need to reread the problem and use another strategy?” and “Does the answer make sense”. Jacobse and Harskamp (2009) used a computer program with problem solving hints in the form of the so called ‘Task stairs’ supporting students’ analysis and exploration of the problem, planning, monitoring and evaluation. The authors used prompts and questions like "what is the question" and "write down the calculation and check your answer carefully" to stimulate students to solve word problems systematically. Much like aforementioned studies, the computer program of Chang, Sung and Lin (2006) named MathCAL, was designed around four problem solving episodes: Understanding the problem, making a plan, executing the plan and reviewing the solution. Lastly, the study of Chung and Tam (2006) adds to the study of Teong. This study showed that cognitive instruction aimed at supporting students with mild intellectual abilities to read a word problem aloud, paraphrase, visualize, state the problem, hypothesize, estimate, calculate, and evaluate can significantly affect the problem solving performance of these students. Providing worked examples of problem solving steps seems particularly useful for mild ability students.

A particular characteristic which all aforementioned studies have in common is that they use some type of visual representation to support problem solving. Apparently, letting students make a sketch or providing them with visualizations, support their comprehension of the underlying structure of the problems which in turn enhances their problem solving performance. The approaches differ as to the scope of the approach (the whole process or just the first and probably hardest part of it) and the structuring of teaching. For students with low mathematics ability direct instruction with the use of worked examples (as used in the research on schema based instruction) may be most effective.

6. Conclusion and Discussion

In this meta-analysis, we have examined the impact of instructional interventions on mathematics learning. A broad goal of this work is to extract important factors that contribute to the improvement of mathematics teaching in schools and we have achieved such goal by identifying the critical "ingredients" of the effective use of instructional interventions. This broad goal certainly limited our effort to the "whether" aspect rather than the "why" aspect of using certain, more effective types of interventions for mathematics learning. The "why" issue requires an in-depth analytical approach and specifically designed research (particularly experimental research) to manipulate key elements in instructional method, use of visual models, use of computer support, grouping of students and students prior skills in mathematics. Few empirical studies we retrieved were specific enough in this regard. Nevertheless, our meta-analysis did imply that the use of a systematic teaching model, that includes systematic sequencing of learning tasks, visualization of the solution steps needed for a tasks, use of strategies to solve problems are in general effective instructional components that promote mathematics learning of students. We do believe that these key ingredients might lead future researchers to investigate the why aspect of more effective mathematics teaching.

Limitations

This meta-analysis has its limitations. First, the fact that meta-analysis cannot be conducted in an experimental fashion exposes our inability to control sample sizes and missing data. We have to acknowledge that only a few study on special education were included in our meta-analysis, so it is not well founded to draw conclusions for special education. Second, small samples lead to the decreased sensitivity of data analysis. Weighting effect sizes by sample sizes restores sensitivity only to a certain extent. Third, it is impossible for any meta-analysis to evaluate (and code) the design quality of the programs (instructional intervention in our case) used in primary studies. In other words, we cannot control the intervention integrity and implementation fidelity. For instance, in most studies there is only scant information of the curriculum in the control group. The results of meta-analysis evaluating the effectiveness of educational interventions are most likely populated without such knowledge. Fourth, we acknowledge the categories for types of instructional methods, types of students, grouping arrangements and duration of the interventions are broad and possible usage can only be instructionally helpful if the primary studies are consulted (we enclosed a list of references).

Overall Effects

Our first research question is: 'Does mathematics learning with explicit instructional interventions impact mathematics achievement of K-6 students, compared to mathematics learning in traditional settings?' We found overall positive effects of new instructional interventions on mathematics achievement. All interventions used an explicit instructional model, to teach students in a sub-domain of mathematics. Of the 40 primary studies measuring effectiveness of mathematical intervention, the average weighted random effect size *d* was .58 with a 95% confidence interval from .45 to .72. This is a moderate but significantly positive effect on mathematics achievement. The effect indicates that students learning mathematics with the new interventions, compared to those following traditional teaching had higher mathematics achievement. This result stresses the importance of systematic teaching according to an empirically validated teaching model. Not all teaching approaches however, resulted in equally good performance and not all students learning with new interventions learned better than those learning without new interventions under all conditions.

Sample and intervention related factors that moderate the effects

The second research question is: What study features moderate the effects of instructional interventions on K-6 students' mathematics achievement? As presented in table 2, we did not find differences between studies using with an experimental design and random allocation of students (23 studies) versus studies with a quasi-experimental studies (17 studies). However, the kind of outcome measure in the studies did significantly affect the average effect size. The average effect size found in the 31 studies using a researcher developed test is .64 (*SE*=.07), while the effect on standardized tests in the 9 remaining studies is significantly lower (*d*=.32, *SE*=.13). This confirms previous findings in meta-analyses of interventions in education (Li & Ma, 2010).

As shown in table 3, we found no significant difference in the effect of instructional interventions for low ability versus high ability students or for students in special education versus students in regular education. This seems in contrast to the findings of the meta-analysis by Kroesbergen en Van Luit (2003). They proposed that with students who lag behind, the learning gap to bridge is wider and the opportunity to make a difference becomes larger. But, in our research the studies with low ability students and students in special education were under-sampled which makes it difficult to show significant differences.

To examine effects of characteristics of the instructional methods in the studies we coded several implementation features: (a) method of teaching and (b) set-up of the instruction and (c) duration of the intervention. Table 4 shows that we found no significant intervention-related characteristics which moderated effects on mathematics achievement. In general, we found no indication that indirect instruction (following the constructivist approach of guiding students instead of leading them) resulted in better student achievement than direct instruction. (However, as we reported earlier, some of the studies in Operations and Ratio, that were specially designed to test the effect of these methods of teaching led to differences in achievement). But, we in our overall analysis we could not investigate (due to the limited number of studies) whether direct instruction is more profitable for special education students or low math ability students.

On average, small group interventions or individual interventions taken together yielded higher effects than whole class interventions. This is often found in meta-analysis (Slavin and Smith, 2008). There was no difference in effect for short-term versus long term interventions (this finding complies with the results of Kroesbergen en Van Luit, 2003, that long term studies do not lead to higher learning gains).

Domain specific Effects and suggestions for further research

The third research question is: What are the optimal conditions for effective mathematics teaching with the interventions in terms of K-6 students' mathematics achievement in key mathematical sub-domains? To get a clear view about what works for different types of mathematical content, we discuss effectiveness and characteristics of interventions specific to the domains of Number Sense, Operations, Fractions, Ratio and Percentages; Measurement and Geometry; and Word Problem Solving. The average effect sizes within the different domains in table 5 show that studies in all different domains report medium to large effects of interventions. Smaller effects were reported in the complex domain of measurement and geometry, while number sense interventions relatively had the largest mean effects.

<u>Number sense</u>

The overall effect size of the interventions on number sense is more than moderate; Cohen's d = .67 (SE=.07). The studies in this sub-domain can be divided into two groups. The first group deals with extensive treatments of number sense. The interventions took place during many weeks and students were trained every day for a brief period in different aspects of number sense. The activities entail: Counting objects and counting on or backwards, comparison of quantities and numbers, change operations (addition and subtraction with quantities) and measurement of length, weight and volume (with natural measures). The activities are brought to children through books, songs, computer games, projects, role play or group discussions. They represent the various ways teachers are used to create opportunities for young children to learn and play. To test the effect of the programs a broad spectrum of subtests is used. The second group of studies uses brief interventions mainly aimed at number identification and comparison. The studies used simple board games or computer games in which children have to count on a number line of squares. The games improved children's number recognition and numerical comparison skills considerably. The effects of the second group of studies are in the topic of number comparison quite comparable to effects of those in the first group of studies. It shows that number sense can be trained in various, and also time-efficient manners. It should be noted that the control groups in the number sense intervention studies typically continued the program kindergartens offered and did not receive systematic education in pre-school mathematics. In comparison, the intervention programs offered a systematic approach to learning early mathematic skills.

Suggestions for further research:

a) What is the added value of number sense programs in kindergarten with a broader scope compared to interventions on number sense restricted to counting and number line practice for the prediction of mathematics achievement in grade 1 and 2?

b) Is the effect of number sense interventions mainly due to less systematic number sense practice in regular kindergarten or is it also due to the use of specific components of the number sense interventions?

Number Operations

The sub-domain of Operations is about addition, subtraction, multiplication and division computation with whole numbers. Studies in the domain of operations on average have a medium effect on mathematics achievement: d=.50 (SE=.05). The intervention studies can be subdivided into two groups: individualized practice studies, and studies on group teaching. The individualized practice studies are about the use of various visualizations for representing numbers and operations (colored bars on a number line, unifix bricks hidden under a cloth, visual supports with blocks and number lines and multiplication grids) and their effect on number operations skills. In these studies visualizations were used to bridge the gap between counting and calculation strategies. In some of the studies the program or teacher determines the use of visuals and the level of tasks and in other studies the student is challenged to deliberately practice with a program and make their own choices. The studies on group teaching are about peer assisted learning strategies or teaching strategies and the use of the number line with concrete materials (frame of beads, multiplication grid) and games to learn the number facts to 20 and operations to 100. Several studies tried to establish the effect of direct versus indirect (guided) instruction in special education. Direct instruction seems to be more favorable to special education student, but more to males than females.

The general theme in the studies discussed above is that conceptual knowledge and skills of students are best supported with the help of visual aids and by training of cognitive strategies. This can be achieved by teaching but also with the help of computer programs that are added to teachers' instruction of number operations.

Suggestions for further research

c) Are visual model(s) necessary to teach number operations after students are introduced to numbers and have attained number sense? If so, which visual model(s) are most effective for addition and subtraction or multiplication and division?

d) In which stage in the process of teaching operations starts guided instruction to be more effective than direct instruction (see the assumption of Hiebert and Grouws, 2007)?

Ratio

The studies in the domain of fractions, ratio and percentages showed to have a more than moderate effect of *d*=.63 (*SE*=.19). In the intervention studies much attention is paid to visual models representing rational numbers in proportion, fractions, percentages and decimal numbers. Fractions are conceptualized with the help of circles, chips, bars and number lines. Decimal numbers (fractions) are depicted with a position schema for the different positions of digits in a decimal number and compared to fractions visualized by fraction circles. Numbers in ratio problems are made visible with the help of part-whole schemes.

Several studies put forward that especially in this sub-domain, tht is often hard to understand for students, it is of great importance to give students a very active role in acquiring basic concepts. In one study this was done by having students discuss the outcome of their first trial of a problem and provide hints on how to find a correct solution. The assumption is that through indirect instruction on the use of visualizations and through discussion of strategies ratio concepts will be formed more firmly than through direct instruction. In this review each approach turned out to be more effective than regular practice in ratio problems.

Suggestions for further research

e) Do cognitve conflict and discussions among students, combined with visual models help students to solve ratio problems more effectively than teacher led group instruction and whole group discussions?

f) Which visualization(s) for fractions, ratio and percentages are most effective to enhance understanding of ratio problems and are most efficient for successful ratio calculation?

Measurement and Geometry

Much research has been done on the development of children's understanding of measurement but few experiments have been undertaken. The average effect size of the five studies we found is of medium level (*d*=.48, *SE*=.13). The research is diverse and ranging from research on conceptual knowledge of area and calculation of area with the help of squares and a grid, to the learning of the properties of geometrical figures with the use of computer programs. Digital Tangram puzzles are used to cover figures with different geometrical shapes and there were computer programs for manipulating of geometrical shapes in order to discover the properties of the shapes and comparing their area. And, there is a study in which virtual manipulatives for 3D and 2D figures were used in tool software. Students had to draw shapes, state a rule for a pattern of shapes, use symmetry in shapes and show equal parts in shapes. There is also research on learning of 3D geometric solids (cubes, pyramids, cones and cylinders) with hands-on activities such as: constructing pyramids with cubes, wrapping cones, measuring the shadows of cylinders, etc..

In the five studies we found that for area measurement and geometry, visualizations (such as: squares, grids or virtual shapes) can play an important role. They can be used to bridge the gap between counting units or calculating the area of plane figures and to bridge the gap between concrete manipulation of geometrical shapes and understanding their characteristics. It is of importance that students have opportunities to use these visualization tools for themselves.

Suggestions for further research

g) A longitudinal experimental study is needed from the primary grades to grade 6 into the development of linear, area and volume measurement with the help of a theoretically consistent set of visualizations. It would be interesting to find out if physical manipulatives are less effective than virtual manipulatives (as suggested by Steen & Brooks (2006). Perhaps virtual manipulatives are more flexible to use and make recognizing shapes and manipulating them easier.

h) Especially in the field of measurement it is of great interest to find out if the combination of teaching measurement concepts and calculation in the higher grades is more effective than subsequently teaching of measurement concepts and calculation (as suggested by Huang & Witz, 2011

<u>Word problem Solving</u>

The experimental studies found in our literature search show that problem solving interventions can have a substantial impact on students' word problem solving performance (*d*=.50, *SE*=.12). All studies in this sub-domain try to teach students problem solving steps like the ones Schoenfeld (1992) described. But there are clear differences in the focus of the studies. About half of the studies focus on the first phases of the process: analyzing the problem and constructing a mental model or schema of the problem situation. Students are taught how to recognize and analyze different types of problem situations and find a formula to solve the problem. The studies are often restricted to a specific topic and do not aim at a general approach for solving word problems. The other set of studies apply a broader set of problem solving strategies to enhance problem solving performance. Mostly, a general problem solving approach was offered to students. The interventions are not restricted to one sub-domain but aim at the problem solving process as such.

A particular characteristic which all the aforementioned studies have in common is that they use visual representations to support problem solving. Apparently, letting students make a sketch or providing them with visualizations, supports their comprehension of the underlying structure of the problems which in turn enhances their problem solving performance. The approaches differ as to the scope of the approach (the whole process or just the first and probably hardest part of it) and the structuring of teaching. Especially for students with a mild ability direct instruction with worked examples may be most effective.

Suggestions for further research

i) Much research has been undertaken in the field of word problem solving. Time now to compare the effects of two approaches: a schema-based approach aiming at the analysis of problem situations in specific domains and a more general approach not restricted to sub-domains. An extensive experimental duration will be needed to test the effectiveness of both approaches over sub-domains and student populations (e.g. students in regular and special education). The outcome would be of great interest in view of current division in research program. In the proposed research prior mathematical knowledge and reading skills of students should be taken into account.

j) As the effective use of schemes or more general strategies is highly dependent on metacogntive monitoring and skills of students, it would be advisable to include measures of metacognition in the studies of word problem solving (Jacobse & Harskamp, 2009)

Practical Implications

It is encouraging to report that using new interventions in school settings may improve teachers practice. In our analysis certain approaches to teaching showed larger effects on mathematics achievement than other approaches. This implies that new interventions may work better in a certain type of learning environment. To find out which approach works best for which learning environment, new research is need in which the implementation of the interventions by the teacher plays an important role. In most studies either the experimenter played the role of teacher or the teachers were overseen en monitored by a research assistant who watched over high fidelity implementation of the intervention. It needs to be studied how interventions can be implemented through forms of consultation in which much responsibility of the implementation is left to the teachers. To achieve this, the teacher has to fit in new interventions into the existing classroom practice in which a curriculum is already in use. The teacher has to find out which parts of the old curriculum can be replaced by the new intervention. This may ask for design research and co-operation between researchers and teachers.

References

Battista, M. T. (2003). Understanding students' thinking about area and volume measurement. In D. H. Clements and G. Bright (Eds.), *Learning and teaching measurement* (pp. 122 - 142). Reston, VA: National Council of Teachers of Mathematics.

Bonotto, About students' understanding and learning of the concept of surface area. In: D.H. Clements and G. Bright, Editors, *Learning and teaching measurement. 2003 Year book*, NCTM, Reston, VA (2003), pp. 157–167.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex, England: John Wiley & Sons.

Borenstein M, Hedges L, Higgins J, Rothstein H. (2005). *Comprehensive Meta-analysis Version 2*. Englewood NJ, USA: Biostat.

Chazan, D. & Lehrer, R. (1998) *Designing Learning Environments for Developing Understanding of Geometry and Space*. Lawrence Erlbaum Associates Mahwah, NJ.

Duval, R.: 2000, 'Basic issues for research in mathematics education', *Proceedings of the 24th Conference of the International Group for the Psychology of Mathematics Education. (PME)*, Vol. 1, Hiroshima, Japan, pp. 55–69.

Expert group Mathematics education (2008) *Crossing thresholds in mathematics* (Over de drempels met rekenen). SLO, Enschede, The Netherlands.

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P.M., Capizzi, A. M., et al. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*, 98, 29–43.

Fuchs, L. S., Fuchs, D., Finelli, R., Courey, S. J., & Hamlett, C. L. (2004). Expanding schema-based transfer instruction to help third graders solve real-life mathematical problems. *American Educational Research Journal, 41*(2), 419-445.

Fuchs, L. S., Fuchs, D., Stuebing, K., Fletcher, J. M., Hamlett, C. L., & Lambert, W. (2008). Problem solving and computational skill: Are they shared or distinct aspects of mathematical cognition? *Journal of Educational Psychology, 100*(1), 30-47.

Hedges, L. and I. Olkin (1985). *Statistical methods for meta-analyses*. Orlando, FL, Academic Press.

Heuvel-Panhuizen, M. (1996). *Assessment and realistic mathematics education*. Dissertation. University of Utrecht, Utrecht.

Hiebert, J. en D.A. Grouws (2007).'The effects of classroom mathematics teaching on tudents' learning.' In F. K. Lester (ed.), S*econd handbook of research on mathematics teaching and learning,* 371-404. Greenwich, CT: information Age Publishing.

Iszak, A. (2005) You have to count the squares: Applying knowledge in pieces to learning rectangular area. *The Journal of Learning Sciences*, 14 , 3, 361-403.

Jordan, N. C., Kaplan, D., Ol áh, A. N., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development*, 77 (1), 153-175.

Kamii, C., & Kysh, J. (2006). The difficulty of "Length x width": Is a square the unit of measurement? *The Journal of Mathematical Behavior, 25*(2), 105-115.

Kroesbergen, E. H., & van Luit, J. E. H. (2003). Mathematics interventions for children with special educational needs: A meta-analysis. *Remedial and Special Education, 24*(2), 97-114.

Kroesbergen, E. H., & van Luit, Johannes E. H. (2002). Teaching multiplication to low math performers: Guided versus structured instruction. *Instructional Science, 30*(5), 361-78.

Li, Q., & Ma, X. (2010). A meta-analysis of the effects of computer technology on school students' mathematics learning. *Educational Psychology Review, 22*(3), 215-243.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA US: Sage Publications, Inc.

Mayer, R. E., & Hegarty, M. (1996). The process of understanding mathematical problems. In R. J. Sternberg, T. Ben-Zeev, R. J. Sternberg & T. Ben-Zeev (Eds.), *The nature of mathematical thinking* (pp. 29-53). Hillsdale, NJ England: Lawrence Erlbaum Associates, Inc.

Meelissen, M.R.M. & Drent, M. (2008) *TIMMS-2007 The Netherlands*, University of Twente, The Netherlands.

Ministry of Education (2010) *Implementation of standards for Dutch language and mathematics* (Invoering referentieniveaus Taal en Rekenen). Den Haag: Netherlands Ministry of Education.

Mullis, I.V.S., Martin, M.O., & Foy, P. (2008). *TIMSS 2007 International Mathematics Report. Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eigth Grades.* Boston: Boston College, TIMSS & PIRLS International Study Center.

National Council of Teachers of Mathematics [NCTM] (2000). *Principles and Standards for School Mathematics.* Reston, VA.,

Rosenthal, R. and R. L. Rosnow (1984). *Essentials of behavioral research: Methods and data analysis*. New York, McGraw-Hill.

OECD (2010), *PISA 2009 at a Glance*, OECD Publishing.

Schoenfeld, A. H. (1992). Learning to think mathematically: problem solving, metacognition, and sense making in mathematics. In D.A.Grouws (Ed.), *Handbook of research on mathematics teaching.* (pp. 224-270). New York: McMilan Publishing.

Slavin, R. & Smith, D. (2008) Effective Programs in Elementary Mathematics: A Best-Evidence Synthesis. *Review of Educational Research*, 78, 3, 477 – 515.

Swanson, H. L., Carson, C., & Sachse-Lee, C. (1996). A selective synthesis of intervention research for students with learning disabilities. *School Psychology Review, 25*(3), 370-91. Tornroos, J., 2005, Mathematics textbooks, opportunity to learn and student achievement, *Studies in Educational Evaluation* 31 (2005) 315-327.

US Department of Education (2008) *Foundations for Success: The Final Report of the NationalMathematics Advisory Panel.* Washington: US Department of Education.

van Garderen, D. (2004). Reciprocal teaching as a comprehension strategy for understanding mathematical word problems. *Reading and Writing Quarterly*, *20*, 225-229.

Verschaffel, L., De Corte, E., Lasure, S., Van Vaerenbergh, G., Bogaerts, H., & Ratinckx, E. (1999). Learning to solve mathematical application problems: A design experiment with fifth graders. *Mathematical Thinking and Learning, 1*(3), 195-229.

References of Studies Used in The Meta-Analysis

Arnold, D. H., Fisher, P. H., Doctoroff, G. L., & Dobbs, J. (2002). Accelerating math development in head start classrooms. *Journal of Educational Psychology, 94*(4), 762-770.

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development, 79*(4), 1016-1031.

Chang, K., Sung, Y., & Lin, S. (2006). Computer-assisted learning for mathematical problem solving. *Computers and Education, 46*(2), 140-151.

Chung, K. K. H., & Tam, Y. H. (2005). Effects of cognitive-based instruction on mathematical problem solving by learners with mild intellectual disabilities. *Journal of Intellectual and Developmental Disability, 30*(4), 207-216.

Clements, D. H., & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the building blocks project. *Journal for Research in Mathematics Education, 38*(2), 136-163.

Clements, D. H., & Sarama, J. (2008). Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Educational Research Journal, 45*(2), 443-494.

Cramer, K. A., Post, T. R., & delMas, R. C. (2002). Initial fraction learning by fourth- and fifth-grade students: A comparison of the effects of using commercial curricula with the effects of using the rational number project curriculum. *Journal for Research in Mathematics Education, 33*(2), 111-144.

Desoete, A., Roeyers, H., & De Clercq, A. (2003). Can offline metacognition enhance mathematical problem solving? *Journal of Educational Psychology, 95*(1), 188-200.

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., et al. (2010). The effects of strategic counting instruction, with and without deliberate practice, on number combination skill among students with mathematics difficulties. *Learning and Individual Differences, 20*(2), 89-100.

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., et al. (2009). Remediating number combination and word problem deficits among students with mathematics difficulties: A randomized control trial. *Journal of Educational Psychology, 101*(3), 561-576.

Fuchs, L. S., Fuchs, D., Yazdian, L., & Powell, S. R. (2002). Enhancing first-grade children's mathematical development with peer-assisted learning strategies. *School Psychology Review, 31*(4), 569-83.

Fuchs, L. S., Seethaler, P. M., Powell, S. R., Hamlett, C. L., & Fletcher, J. M. (2008). Effects of preventative tutoring on the mathematical problem solving of third-grade students with math and reading difficulties. *Exceptional Children, 74*(2), 155-173.

Fuchs, L. S., Zumeta, R. O., Schumacher, R. F., Powell, S. R., Seethaler, P. M., Hamlett, C. L., et al. (2010). The effects of schema-broadening instruction on second graders' word-problem performance and their ability to represent word problems with algebraic equations: A randomized control study. *Elementary School Journal, 110*(4), 440-463.

Hannafin, R. D., Truxaw, M. P., Vermillion, J. R., & Liu, Y. (2008). Effects of spatial ability and instructional program on geometry achievement. *Journal of Educational Research, 101*(3), 148-157.

Hohn, R. L., & Frey, B. (2002). Heuristic training and performance in elementary mathematical problem solving. *Journal of Educational Research, 95*(6), 374-80.

Huang, T., Liu, Y., & Shiu, C. (2008). Construction of an online learning system for decimal numbers through the use of cognitive conflict strategy. *Computers & Education, 50*(1), 61-76.

Huang, H. E., & Witz, K. G. (2011). Developing children's conceptual understanding of area measurement: A curriculum and teaching experiment. *Learning and Instruction, 21*(1), 1-13.

Jacobse, A. E., & Harskamp, E. G. (2009). Student-controlled metacognitive training for solving word problems in primary school mathematics. *Educational Research and Evaluation, 15*(5), 447-463.

Kroesbergen, E. H., & van Luit, J. E. H. (2002). Teaching multiplication to low math performers: Guided versus structured instruction. *Instructional Science, 30*(5), 361-78.

Kroesbergen, E. H., Van Luit, J. E. H., & Maas, C. J. M. (2004). Effectiveness of explicit and constructivist mathematics instruction for low-achieving students in the Netherlands. *Elementary School Journal, 104*(3), 233.

Nunes, T., Bryant, P., Hallett, D., Bell, D., & Evans, D. (2009). Teaching children about the inverse relation between addition and subtraction. *Mathematical Thinking and Learning: An International Journal, 11*(1-2), 61-78.

Olkun, S., Altun, A., & Smith, G. (2005). Computers and 2D geometric learning of Turkish fourth and fifth graders. *British Journal of Educational Technology, 36*(2), 317-326.

Ploger, D., & Hecht, S. (2009). Enhancing children's conceptual understanding of mathematics through chartworld software. *Journal of Research in Childhood Education, 23*(3), 267.

Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development, 79*(2), 375-394.

Schoppek, W., & Tulis, M. (2010). Enhancing arithmetic and word-problem solving skills efficiently by individualized computer-assisted practice. *Journal of Educational Research, 103*(4), 239-252.

Siegler, R. S., & Ramani, G. B. (2008). Playing linear numerical board games promotes low-income children's numerical development. *Developmental Science, 11*(5), 655-661.

Souvignier, E., & Kronenberger, J. (2007). Cooperative learning in third graders' jigsaw groups for mathematics and science with and without questioning training. *British Journal of Educational Psychology, 77*(4), 755-771.

Starkey, P., Klein, A., & Wakeley, A. (2004). Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention. *Early Childhood Research Quarterly, 19*(1), 99-120.

Steen, K., Brooks, D., & Lyon, T. (2006). The impact of virtual manipulatives on first grade geometry instruction and learning. *Journal of Computers in Mathematics and Science Teaching, 25*(4), 373-391.

Tajika, H., Nakatsu, N., Nozaki, H., Neumann, E., & Maruno, S. (2007). Effects of self-explanation as a metacognitive strategy for solving mathematical word problems. *Japanese Psychological Research, 49*(3), 222-233.

Teong, S. K. (2003). The effect of metacognitive training on mathematical word-problem solving. *Journal of Computer Assisted Learning, 19*(1), 46-55.

Timmermans, R. E., Van Lieshout, Ernest C. D. M., & Verhoeven, L. (2007). Gender-related effects of contemporary math instruction for low performers on problem-solving behavior. *Learning and Instruction, 17*(1), 42-54.

Tournaki, N. (2003). The differential effects of teaching addition through strategy instruction versus drill and practice to students with and without learning disabilities. *Journal of Learning Disabilities, 36*(5), 449-58.

Van Dijk, I. M. A. W., Van Oers, B., Terwel, J., (2003). Strategic learning in primary mathematics education: Effects of an experimental program in modelling. *Educational Research and Evaluation: An International Journal on Theory and Practice, 9*(2), 161-87.

Van Luit, Johannes E. H., & Schopman, E. A. (2000). Improving early numeracy of young children with special education needs. *Remedial and Special Education, 21*(1), 27-40.

Wilson, A. J., Dehaene, S., Dubois, O., & Fayol, M. (2009). Effects of an adaptive game intervention on accessing number sense in low-socioeconomic-status kindergarten children. *Mind, Brain, and Education, 3*(4), 224-234.

Wong, M., & Evans, D. (2007). Improving basic multiplication fact recall for primary school students. *Mathematics Education Research Journal, 19*(1), 89-106.

Woodward, J. (2006). Developing automaticity in multiplication facts: Integrating strategy instruction with timed practice drills. *Learning Disability Quarterly, 29*(4), 269-289.